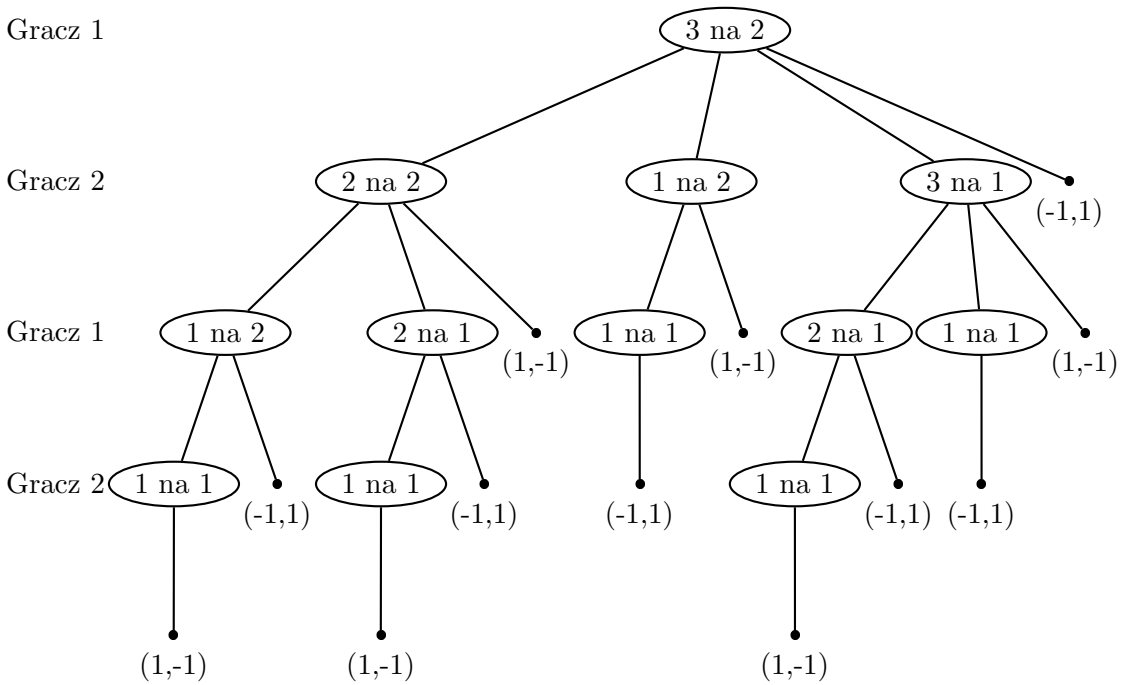
Rozdział 2

Gry niekooperacyjne w postaci ekstensywnej

Wszystkie gry, które dotychczas rozważaliśmy, miały tylko jeden etap. W związku z tym musieliśmy zakładać, że gracze podejmują swoje decyzje równocześnie, nie znając tych, które podjął przeciwnik. Oczywiście w rzeczywistości rzadko się zdarzają takie sytuacje, że decyzje podejmuje się naraz, bez żadnej wiedzy o ruchach przeciwnika. W praktyce prawopodobnie musiałoby to oznaczać, że gracze, jak w dylematie więźnia są od siebie odizolowani. W związku z tym powstał pomysł, żeby kolejne, wykonywane po sobie ruchy opisać przy pomocy drzewa. Na przykład spróbujmy sobie opisać taką grę: Mamy tabliczkę czekolady 3×2. W każdym kolejnym ruchu (gracze grają na przemian) odłamuje się kawałek czekolady, ale wolno odłamywać tylko całe „wiersze" lub „kolumny" czekolady. Ten, kto zje ostatni kawałek – przegrywa. Taką grę można opisać przy pomocy następującego drzewa:



Ogólnie tego typu grę będzie można zdefiniować następująco:

Definicja 2.1 Gra w postaci ekstensywnej (pozycyjna) N graczy opisana jest za pomocą drzewa uporządkowanego T oraz:

1. partycji $P^0, P^1, \ldots, P^N$ wierzchołków drzewa (pozycji), nie będących liściami. P^i, $i \neq 0$ interpretujemy jako zbiór pozycji i-tego gracza (tzn. pozycji, w których on

podejmuje decyzję). P^0 jest zbiorem pozycji „natury" (dalsza pozycja wybierana jest przez mechanizm losowy);

2. dla każdego gracza i, partycji P^i na k^i zbiorów informacyjnych $U^i_1, \ldots, U^i_{k^i}$, takich że:

   (a) z każdego wierzchołka (pozycji) $v \in U^i_j$ wychodzi tyle samo łuków,

   (b) każda droga z pozycji początkowej (korzenia drzewa) do pozycji końcowej (liścia) przecina każdy zbór informacyjny co najwyżej jeden raz;

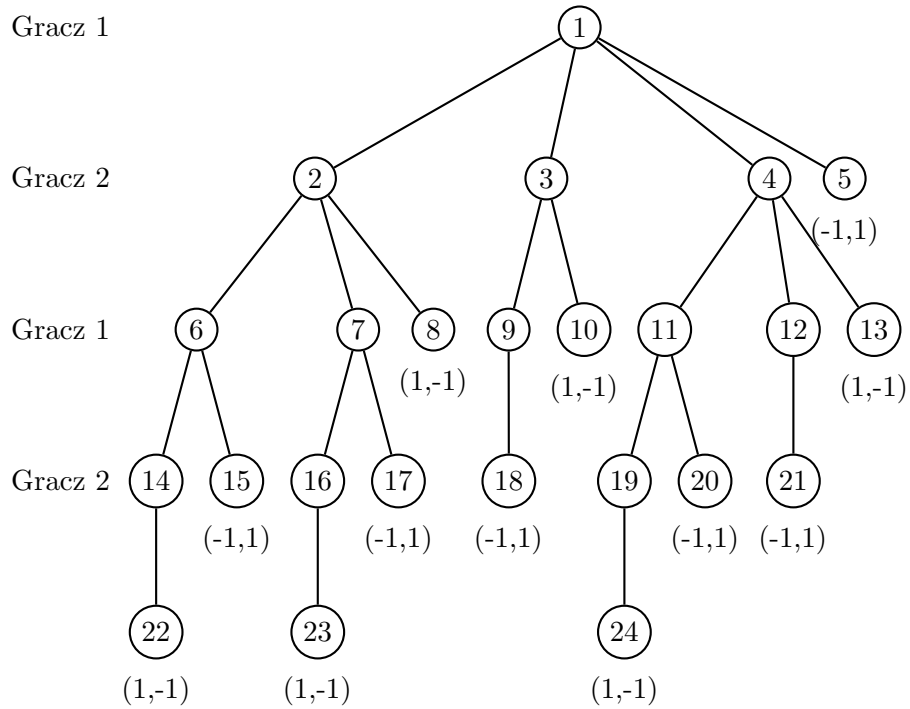
   Interpretacja zbiorów informacyjnych jest taka, że podział na zbiory informacyjne odpowiada stanowi wiedzy gracza – znajdując się w pozycji v, wie tylko tyle, że znajduje się w zbiorze informacyjnym, do którego v należy, i opierając się na tej wiedzy podejmuje decyzje co do swojego działania.

3. dla każdej pozycji natury $v \in P^0$ dany jest rozkład prawdopodobieństwa p_v, określający prawdopodobieństwo, z jakim zostaną wylosowane poszczególne łuki wychodzące z v.

4. dla każdej pozycji końcowej (liścia) t dany jest wektor $g(t) = (g^1(t), \ldots, g^N(t))$ wypłat poszczególnych graczy.

Drogę łączącą korzeń z dowolnym liściem nazywamy partią gry.

Nasz przykład możemy opisać zgodnie z tą definicją następująco: Zacznijmy od ponumerowania wierzchołków naszego drzewa (żeby wszystko było czytelne, tym razem rysuję drzewo bez opisu, co który wierzchołek oznacza – pozostaje jedynie numeracja.



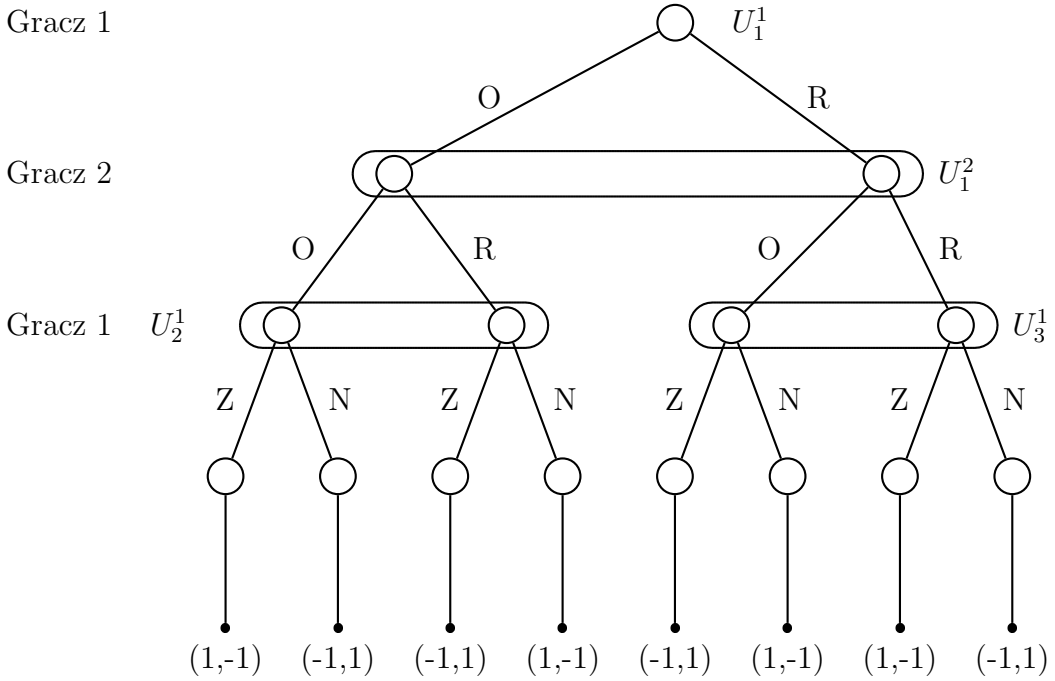
Pozycjami pierwszego gracza będą wierzchołki $P^1 = \{1, 6, 7, 9, 11, 12\}$, pozycjami drugiego – $P^2 = \{2, 3, 4, 14, 16, 19\}$.
Wszystkie zbiory informacyjne są jednoelementowe, bo w każdej chwili gracz zna całą dotychczasową rozgrywkę: $U^1_1 = \{1\}$, $U^1_2 = \{6\}$, $U^1_3 = \{7\}$,..., $U^2_1 = \{2\}$, $U^2_2 = \{3\}$, $U^2_3 = \{4\}$,...
Wypłaty graczy, ponieważ w tej grze chodzi tylko o wygraną bądź przegraną, są zawsze

równe 1 lub -1, a dokładniej: $g(5) = (-1, 1)$, $g(8) = g(10) = g(13) = (1, -1)$, $g(15) = g(17) = g(18) = g(20) = g(21) = (-1, 1)$, $g(22) = g(23) = g(24) = (1, -1)$.

Ale oczywiście definicja gry pozycyjnej jest bardziej złożona, dzięki czemu obejmuje ona też różne bardziej skomplikowane gry.

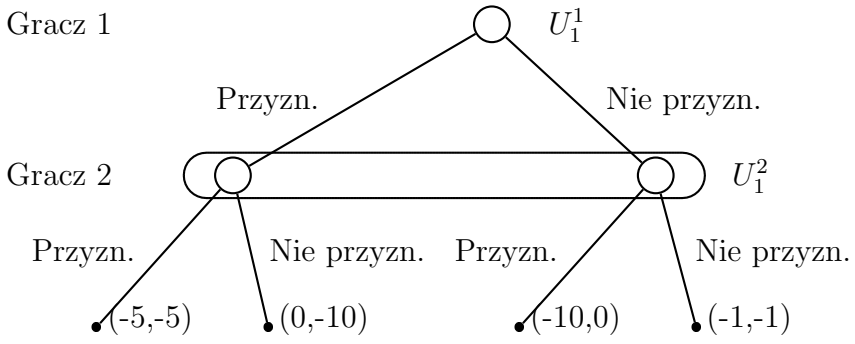
Przykład: Pierwszy gracz rzuca monetą, drugi nie zna wyniku tego rzutu i zgaduje, jaki był, drugi wie, jaki był wynik losowania, i zgaduje, czy 1. odgadł ten wynik. Jeśli zgadnie, to wygrywa, jeśli nie – przegrywa.

Grę będzie opisywało następujące drzewo (zbiory informacyjne są od razu zaznaczone na drzewie):



Tu już widać sens wprowadzenia czegoś takiego jak zbiory informacyjne – w tym a tym przypadku chcemy założyć, że gracz nie zna jakichś ruchów przeciwnika lub natury, i do tego są one nam potrzebne. Ale okazuje się, że stosując taki model możemy opisać coś bardziej skomplikowanego – mianowicie ograniczoną pamięć – gracz może nie pamiętać jakiegoś własnego wyboru z przeszłości. Przykład takiej gry pojawi się na ćwiczeniach. Ten przykład może wydać się trochę abstrakcyjny (bo jak można zapomnieć, co się robiło w pierwszym ruchu, ruch później?), natomiast sama idea tego, że ruchy można zapominać jest w sumie dosyć, „życiowa" (tak się dzieje na przykład w przypadku gier z dużą liczbą etapów).

Zauważmy także, że gry w postaci strategicznej też możemy modelować przy pomocy tego. Np. dylemat więźnia będzie opisywało następujące drzewo (zaznaczone są na nim także zbiory informacyjne)



Dobrze, skoro wiemy, jak opisać grę wieloetapową, to teraz zastanówmy się, jak opisać strategie w takiej grze. Mówiłem, że gracz znajdujący się w pozycji v wie tylko, w jakim zbiorze informacyjnym się znajduje. W związku z tym strategię gracza i będziemy definiować następująco:

Niech $I^i = \{U^i_1, \ldots, U^i_{k^i}\}$ – rodzina zbiorów informacyjnych gracza i, a $\nu(U^i_j)$ – liczba łuków wychodzących z każdej pozycji w zbiorze U^i_j. **Strategią czystą** gracza i nazwiemy odwzorowanie $s^i : I^i \to \mathbb{N}$ takie że $s^i(U^i_j) \leqslant \nu(U^i_j)$. Jeśli gracz znajduje się w pozycji $v \in U^i_j$ i stosuje strategię s^i, to wybiera w tej pozycji łuk wychodzący z v o „numerze" $s^i(U^i_j)$.

Wypłatą gracza i, gdy poszczególni gracze stosują strategie $s = (s^1, \ldots, s^N)$ jest $\sum_{v \in L(T)} g(v)p_s(v)$, gdzie $L(T)$ oznacza zbiór liści drzewa, a $p_s(v)$ jest prawdopodobieństwem dojścia do liścia v, jeśli stosowane są strategie s.

Zauważmy jednak, że jesłi zbiory I^i są skończone (a tak zakładamy), to grę w postaci ekstensywnej możemy zawsze sprowadzić do gry w postaci strategicznej, za strategie w grze jednokrokowej przyjmując ciągi strategii w kolejnych zbiorach informacyjnych. Na przykład w przedostatnim przykładzie (z rzutem monetą), pierwszy gracz miał trzy zbiory informacyjne, a w każdym z nich – dwie strategie (w pierwszym O (orzeł) i R (reszka), w pozostałych Z (gracz 2. zgadł, co wypadło) i N (nie zgadł)), stąd w grze pozycyjnej miał 8 strategii (czystych): OZZ, OZN, ONZ, ONN, RZZ, RZN, RNZ, RNN. Z kolei gracz drugi miał dwie strategie: O (orzeł) i R (reszka).

Skoro jednak tak, to znaczy, że na mocy twierdzenia Nasha każda gra pozycyjna ma równowagę w strategiach mieszanych. No właśnie – tylko jak będą tutaj wyglądać strategie mieszane? Jeśli chcemy stosować twierdzenie Nasha, to strategią mieszaną będzie musiał być rozkład na zbiorze strategii czystych, czyli funkcjach z I^i w $\mathbb{N}$ – innymi słowy, chcąc używać strategii mieszanej, musielibyśmy dokonać losowania strategii w całej grze przed rozpoczęciem gry. Jeśli to nie byłaby taka mała, paroetapowa gra, to zbiory, z których musielibyśmy losować byłyby olbrzymie, w dodatku losowanie przed rozpoczęciem rozgrywki kłuci się z intuicją co do tego, na czym polega gra wieloetapowa. W związku z tym wprowadza się mniejszy zbiór strategii mieszanych, nazywanych strategiami postępowania.

Definicja 2.2 Strategią postępowania gracza i w grze pozycyjnej zdefiniowanej wcześniej nazwiemy odwzorowanie, które dowolnemu zbiorowi informacyjnemu $U^i_j \in I^i$ przypisuje pewien rozkład prawdopodobieństwa na zbiorze $\{1, \ldots, \nu(U^i_j)\}$. Używanie takiej strategii polega na tym, że w momencie gdy gra dojdzie do zbioru informacyjnego U^i_j gracza i-tego, dokonuje on losowania swojego kolejnego ruchu, stosując rozkład prawdopodobieństwa przypisany temu zbiorowi przez zadaną strategię.

Oczywiście ten zbiór strategii jest istotnie mniejszy od zbioru wszystkich strategii zrandomizowanych, i równowaga, która istnieje na mocy twierdzenia Nasha, może wyjść poza ten zbiór. Jest na szczęście duża klasa gier, dla których można się ograniczyć do strategii postępowania. Są to gry *z doskonałą pamięcią*, tzn. takie, w których każdy z graczy pamięta wszystkie swoje dotychczasowe decyzje. To, że tak w istocie jest, pokazał Kuhn, a my spróbujemy to, jak i istotność założenia o doskonałej pamięci, mniej lub bardziej ściśle uzasadnić sobie na ćwiczeniach.

Na koniec opowiem o innym wyniku, który jest jeszcze bardziej budujący – mianowicie o pewnej klasie gier pozycyjnych, dla której strategie optymalne nie tylko istnieją, ale są w strategiach czystych, i do tego łatwo jest znaleźć.

Definicja 2.3 *Grą z pełną informacją* nazwiemy grę pozycyjną, w której wszystkie zbiory informacyjne są jednoelementowe.

Twierdzenie 2.1 (Kuhn) *W każdej grze z pełną informacją istnieje równowaga w strategiach czystych.*

(Najpierw przedstawią dowód tego twierdzenia, a potem na przykąldzie zilustruję, o co w nim tak naprawdę chodziło).

Dowód: Indukcja ze względu na liczbę wierzchołków w drzewie, niebędących liściami. Dla 1 wierzchołka oczywiście w grze jest tylko jeden gracz, więc strategia, która daje mu najlepszą wypłatę jest też równowagą Nasha.

Dalej, załóżmy, że każda gra posiadająca $\leqslant m$ wierzchołków ma równowagę w strategiach czystych. Niech nasza gra będzie opisana drzewem o $m+1$ wierzchołkach. Rozważymy dwa przypadki:

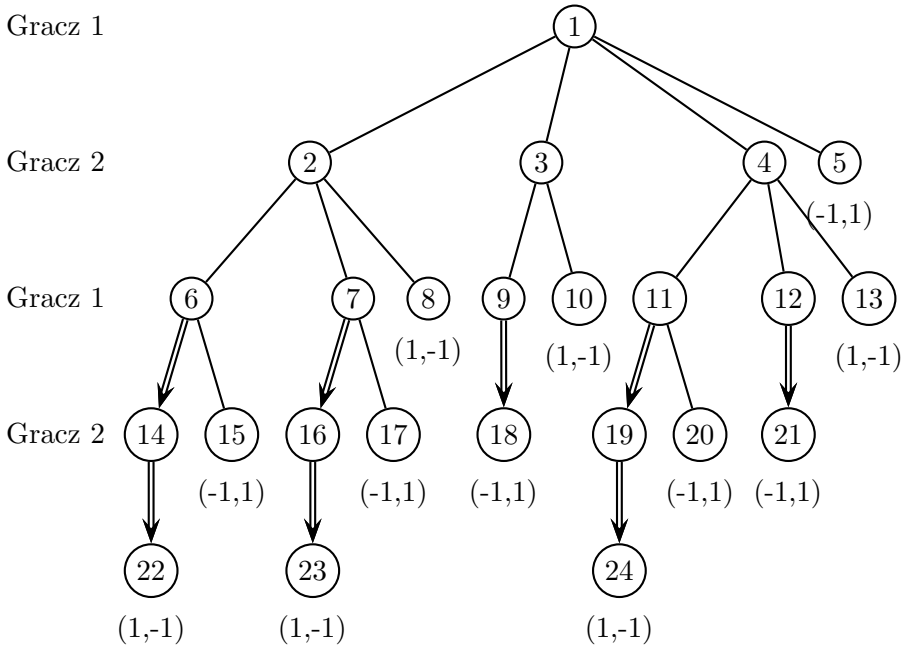
1. Korzeń drzewa r jest pozycją natury. Niech $w_1, \ldots, w_k$ oznaczają pozycje połączone łukiem bezpośrednio z korzeniem, a $T_1, \ldots, T_k$ – drzewa o korzeniach $w_1, \ldots, w_k$. Każde z tych drzew ma $\leqslant m$ wierzchołków, więć gra opisana każdym z tych drzew ma równowagę. Oznaczmy te równowagi przez $s_1, \ldots, s_k$. Wtedy $s(v) = s_j(v)$, jeśli $v \in T_j$ jest równowagą w naszej wyjściowej grze.

2. Korzeń drzewa r jest pozycją gracza i. Oznaczenia, jak poprzednio, dodatkowo niech g_l^i oznacza wypłatę gracza i w równowadze s_l. Wtedy równowagę w grze wyjściowej (tej „dużej") tworzymy następująco. Jeśli $v \in T_i$, to $s(v) = s_i(v)$, natomiast $s(r)$ wybiera tę z krawędzi wychodzących z korzenia l, dla której $g_l^i = \max_{j \leqslant k} g_j^i$.

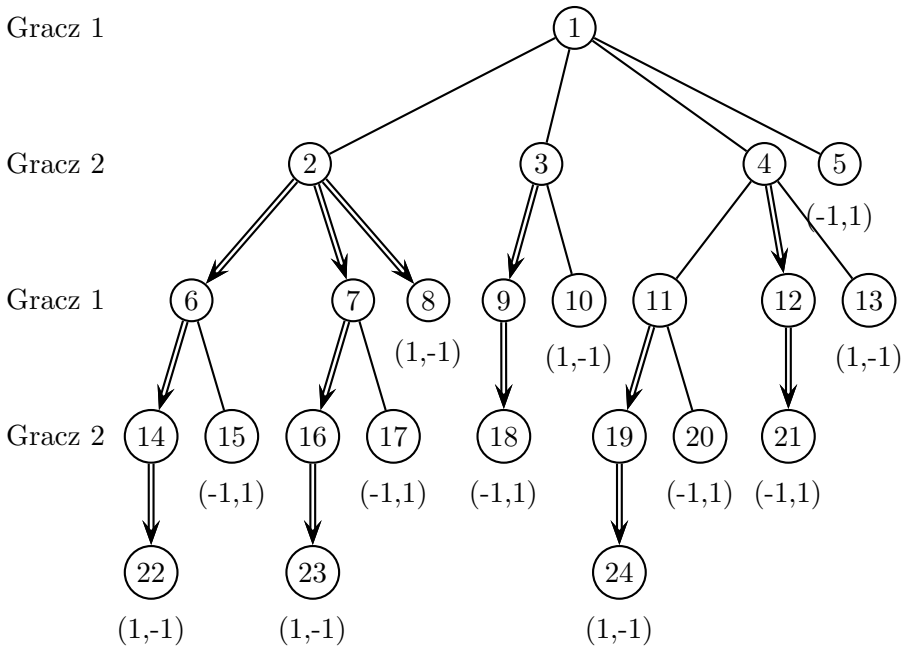
Uzasadnienie, dlaczego tak wybrane strategie tworzą równowagę jest w obu przypadkach takie samo. Załóżmy, że to nie jest równowaga. Wtedy jednemu z graczy opłaca się odstąpić od jej grania. Jeśli nie jest to gracz, którego pozycją był korzeń, to odstępuje od równowagi w pozycji należącej do drzewa T_j. Wtedy jednak musiałoby się to wiązać także z poprawieniem wypłaty w grze z drzewem T_j, co jest sprzeczne z założeniem, że s_j jest równowagą w tej grze. Podobnie w przypadku gracza, którego pozycją jest korzeń – jeśli odstąpił od równowagi właśnie w tej pozycji, to jego wypłata na pewno się zmniejszy (w ten sposób ją zdefiniowaliśmy, żeby się zmniejszyła). Jeśli odstąpi od równowagi w jakiejś innej pozycji, to możemy zastosować rozumowanie, które stosowaliśmy do innych graczy. □

Istotę tego dowodu możemy zilustrować na przykładzie.

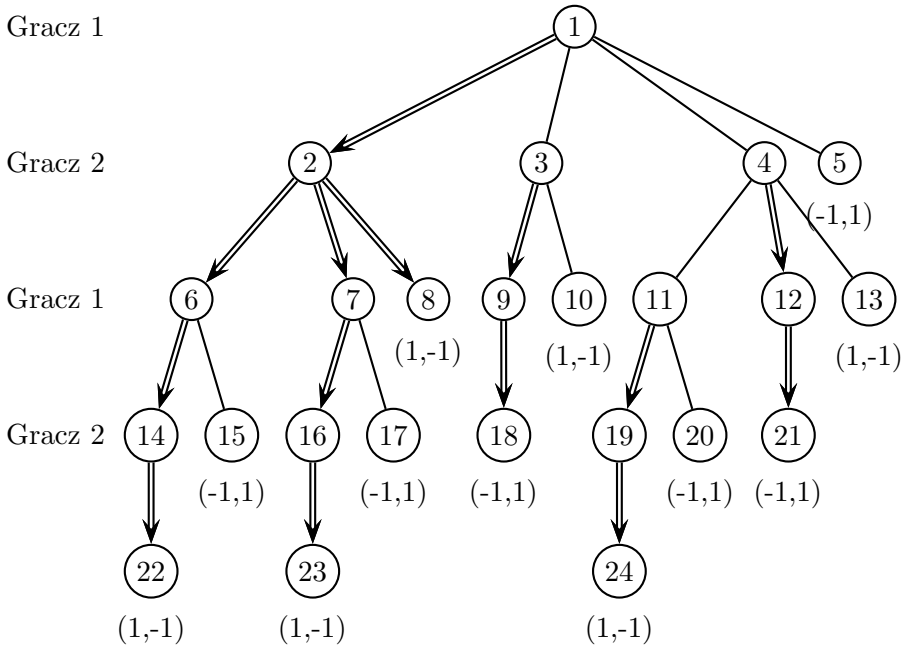
Przykład: Wracamy do gry z początku wykładu. Zaczynamy od samego dołu drzewa. Na 4. etapie Gracz 2. nie ma nic do wyboru. Przechodzimy jeden etap wcześniej. Tutaj Gracz 1. z możliwych zagrań wybiera te, które dają mu większą wypłatę. Na poniższym drzewie są one zaznaczone strzałkami.



Następnie przechodzimy krok wcześniej. Na tym kroku decyzje podejmuje Gracz 2., więc wybiera te krawędzie, dla których jego wypłata jest największa (w przypadku krawędzi wychodzących z wierzchołka 2, zaznaczyłem wszystkie, bo dla Gracza 2. nie ma różnicy, którą z nich wybierze):



No i dochodzimy do pierwszego kroku – na nim decyzję podejmuje Gracz 1., i ma do wyboru wierzchołek 2, w którym, jeśli tylko na dalszych etapach będzie postępował optymalnie (czyli zgodnie ze strzałkami), zapewni sobie wypłatę 1, oraz trzy wierzchołki, w których przeciwnik może sobie zapewnić 1. Wybiera oczywiście 2. wierzchołek:



Uwaga 2.1 Metodę, którą zastosowaliśmy w dowodzie twierdzenia Kuhna (oraz w powyższym przykładzie, nazywamy **metodą indukcji wstecznej**.